

Annexe B - Examen Données Catégorielles : Exploration, Modélisation, 12 décembre 2022

A - Table 1

```
library(FactoMineR)
library(factoextra)
```

```
## Le chargement a nécessité le package : ggplot2
```

```
## Welcome! Want to learn more? See two factoextra-related books at https://goo.gl/ve3WBa
```

```
data.IQS=read.csv2("K:\\myhome\\Documents\\DC\\examen\\2223\\data\\data.txt",row.name=1, header=T)
colnames(data.IQS)<-c("[54.2,91.3[", "[91.3,99.9[", "[99.9,107[", "[107,116[", "[116,156]")
data.IQS
```

##	[54.2,91.3[[91.3,99.9[[99.9,107[[107,116[[116,156]
## AIX-MARSEILLE	301	166	169	251	274
## AMIENS	407	273	224	168	111
## BESANCON	119	172	157	183	87
## BORDEAUX	220	387	408	344	311
## CLERMONT-FERRAND	116	219	236	206	113
## CORSE	50	47	41	19	16
## CRETEIL	501	189	185	245	338
## DIJON	177	242	240	207	119
## GRENOBLE	228	252	400	544	527
## LILLE	836	392	283	247	318
## LIMOGES	65	96	120	93	26
## LYON	277	211	264	350	444
## MONTPELLIER	265	252	294	292	186
## NANCY-METZ	311	313	258	224	156
## NANTES	314	541	519	452	363
## NICE	132	96	157	141	174
## NORMANDIE	370	453	387	343	202
## ORLEANS-TOURS	255	293	269	282	205
## PARIS	46	39	39	57	276
## POITIERS	148	295	223	193	116
## REIMS	215	194	150	111	65
## RENNES	230	450	487	440	343
## STRASBOURG	119	113	144	213	199
## TOULOUSE	193	314	388	429	355
## VERSAILLES	335	178	168	264	912

B - Test du chi2

```
test=chisq.test(data.IQS)
test
```

```
##
## Pearson's Chi-squared test
##
## data: data.IQS
## X-squared = 4037, df = 96, p-value < 2.2e-16
```

C - Profils lignes et colonnes

```
pligne=prop.table(as.matrix(data.IQS),1);round(pligne,3)
```

##	[54.2,91.3[[91.3,99.9[[99.9,107[[107,116[[116,156]
## AIX-MARSEILLE	0.259	0.143	0.146	0.216	0.236
## AMIENS	0.344	0.231	0.189	0.142	0.094
## BESANCON	0.166	0.240	0.219	0.255	0.121
## BORDEAUX	0.132	0.232	0.244	0.206	0.186
## CLERMONT-FERRAND	0.130	0.246	0.265	0.231	0.127
## CORSE	0.289	0.272	0.237	0.110	0.092
## CRETEIL	0.344	0.130	0.127	0.168	0.232
## DIJON	0.180	0.246	0.244	0.210	0.121
## GRENOBLE	0.117	0.129	0.205	0.279	0.270
## LILLE	0.403	0.189	0.136	0.119	0.153
## LIMOGES	0.162	0.240	0.300	0.232	0.065
## LYON	0.179	0.136	0.171	0.226	0.287
## MONTPELLIER	0.206	0.196	0.228	0.227	0.144
## NANCY-METZ	0.246	0.248	0.204	0.177	0.124
## NANTES	0.143	0.247	0.237	0.206	0.166
## NICE	0.189	0.137	0.224	0.201	0.249
## NORMANDIE	0.211	0.258	0.221	0.195	0.115
## ORLEANS-TOURS	0.196	0.225	0.206	0.216	0.157
## PARIS	0.101	0.085	0.085	0.125	0.604
## POITIERS	0.152	0.303	0.229	0.198	0.119
## REIMS	0.293	0.264	0.204	0.151	0.088
## RENNES	0.118	0.231	0.250	0.226	0.176
## STRASBOURG	0.151	0.143	0.183	0.270	0.253
## TOULOUSE	0.115	0.187	0.231	0.256	0.211
## VERSAILLES	0.180	0.096	0.090	0.142	0.491

```
pcol=prop.table(as.matrix(data.IQS),2);round(pcol,3)
```

##	[54.2,91.3[[91.3,99.9[[99.9,107[[107,116[[116,156]
## AIX-MARSEILLE	0.048	0.027	0.027	0.040	0.044
## AMIENS	0.065	0.044	0.036	0.027	0.018
## BESANCON	0.019	0.028	0.025	0.029	0.014

## BORDEAUX	0.035	0.063	0.066	0.055	0.050
## CLERMONT-FERRAND	0.019	0.035	0.038	0.033	0.018
## CORSE	0.008	0.008	0.007	0.003	0.003
## CRETEIL	0.080	0.031	0.030	0.039	0.054
## DIJON	0.028	0.039	0.039	0.033	0.019
## GRENOBLE	0.037	0.041	0.064	0.086	0.085
## LILLE	0.134	0.063	0.046	0.039	0.051
## LIMOGES	0.010	0.016	0.019	0.015	0.004
## LYON	0.044	0.034	0.043	0.056	0.071
## MONTPELLIER	0.043	0.041	0.047	0.046	0.030
## NANCY-METZ	0.050	0.051	0.042	0.036	0.025
## NANTES	0.050	0.088	0.084	0.072	0.058
## NICE	0.021	0.016	0.025	0.022	0.028
## NORMANDIE	0.059	0.073	0.062	0.054	0.032
## ORLEANS-TOURS	0.041	0.047	0.043	0.045	0.033
## PARIS	0.007	0.006	0.006	0.009	0.044
## POITIERS	0.024	0.048	0.036	0.031	0.019
## REIMS	0.035	0.031	0.024	0.018	0.010
## RENNES	0.037	0.073	0.078	0.070	0.055
## STRASBOURG	0.019	0.018	0.023	0.034	0.032
## TOULOUSE	0.031	0.051	0.062	0.068	0.057
## VERSAILLES	0.054	0.029	0.027	0.042	0.146

D - Contributions au chi2

```
tab1=test$res^2; total.l=apply(test$res^2,1,sum)
tab2=cbind(tab1,total.l);total.c=apply(tab2,2,sum)
tab.contrib=rbind(tab2,total.c)
round(tab.contrib,3)
```

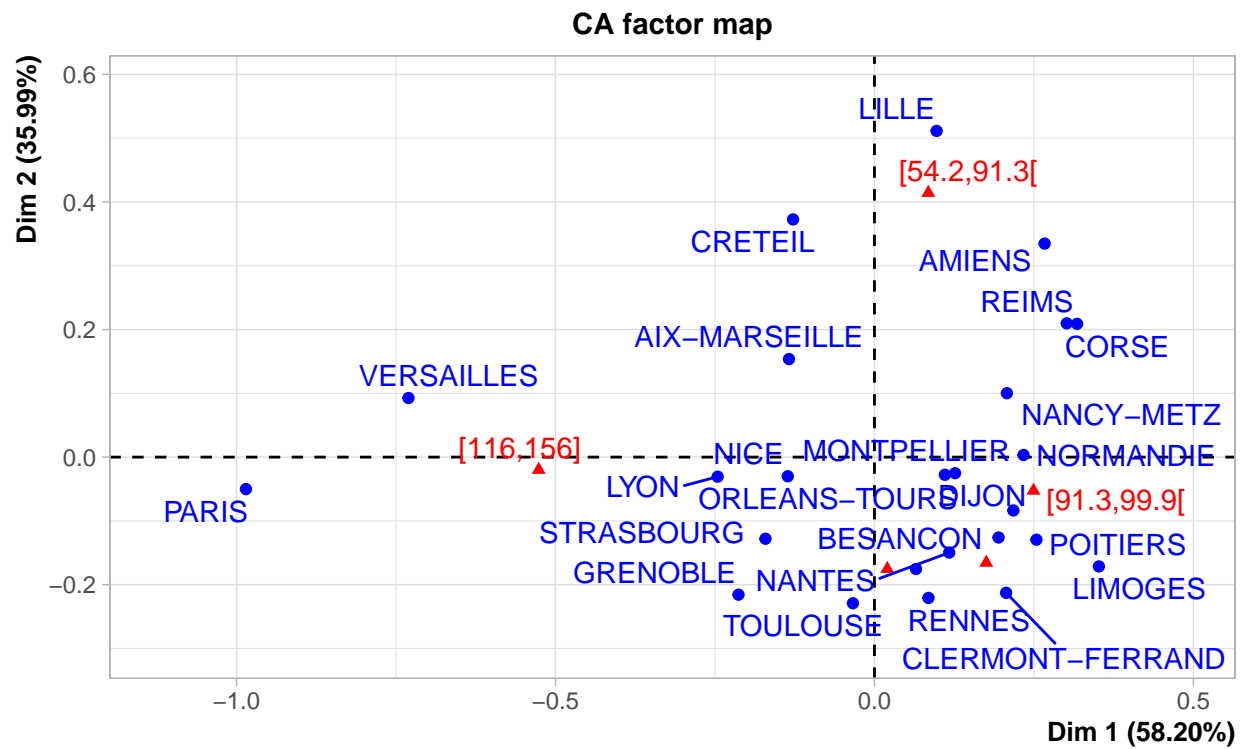
##	[54.2,91.3[[91.3,99.9[[99.9,107[[107,116[[116,156]
## AIX-MARSEILLE	20.390	17.913	16.849	1.128	7.440
## AMIENS	122.737	6.293	0.594	21.181	66.847
## BESANCON	4.213	6.165	1.343	9.862	22.394
## BORDEAUX	38.904	9.420	16.933	0.120	1.625
## CLERMONT-FERRAND	21.592	10.244	19.340	3.775	23.835
## CORSE	6.856	4.698	1.230	7.298	10.024
## CRETEIL	150.390	34.665	38.406	8.404	7.290
## DIJON	2.029	11.157	9.698	0.310	31.000
## GRENOBLE	67.416	47.017	0.315	56.703	47.662
## LILLE	426.516	0.938	41.374	71.075	22.915
## LIMOGES	2.812	3.509	20.326	1.819	36.517
## LYON	3.351	29.788	6.338	4.484	58.463
## MONTPELLIER	0.202	0.051	5.338	3.782	20.112
## NANCY-METZ	13.609	15.737	0.164	3.802	36.964
## NANTES	35.001	26.346	15.642	0.201	12.908
## NICE	0.457	13.200	2.183	0.002	8.186
## NORMANDIE	1.030	31.679	3.942	0.394	63.469
## ORLEANS-TOURS	0.129	4.584	0.315	1.279	12.032
## PARIS	22.549	29.404	29.799	13.559	372.145
## POITIERS	11.326	53.461	4.218	0.086	32.122

## REIMS	31.461	15.977	0.082	9.514	45.852
## RENNES	65.633	10.372	24.839	5.310	5.747
## STRASBOURG	9.452	11.974	1.091	18.090	10.788
## TOULOUSE	60.719	1.077	8.484	23.622	1.061
## VERSAILLES	3.565	98.273	110.437	33.079	785.146
## total.c	1122.338	493.942	379.279	298.880	1742.544
##	total.l				
## AIX-MARSEILLE	63.720				
## AMIENS	217.651				
## BESANCON	43.977				
## BORDEAUX	67.003				
## CLERMONT-FERRAND	78.786				
## CORSE	30.105				
## CRETEIL	239.156				
## DIJON	54.194				
## GRENOBLE	219.112				
## LILLE	562.819				
## LIMOGES	64.983				
## LYON	102.423				
## MONTPELLIER	29.484				
## NANCY-METZ	70.277				
## NANTES	90.098				
## NICE	24.028				
## NORMANDIE	100.514				
## ORLEANS-TOURS	18.338				
## PARIS	467.456				
## POITIERS	101.213				
## REIMS	102.886				
## RENNES	111.902				
## STRASBOURG	51.395				
## TOULOUSE	94.962				
## VERSAILLES	1030.500				
## total.c	4036.983				

E - AFC et graphique du premier plan factoriel

```
AFC=CA(data.IQS)
```

```
## Warning: ggrepel: 3 unlabeled data points (too many overlaps). Consider
## increasing max.overlaps
```



F - Inertie totale et inertie des axes

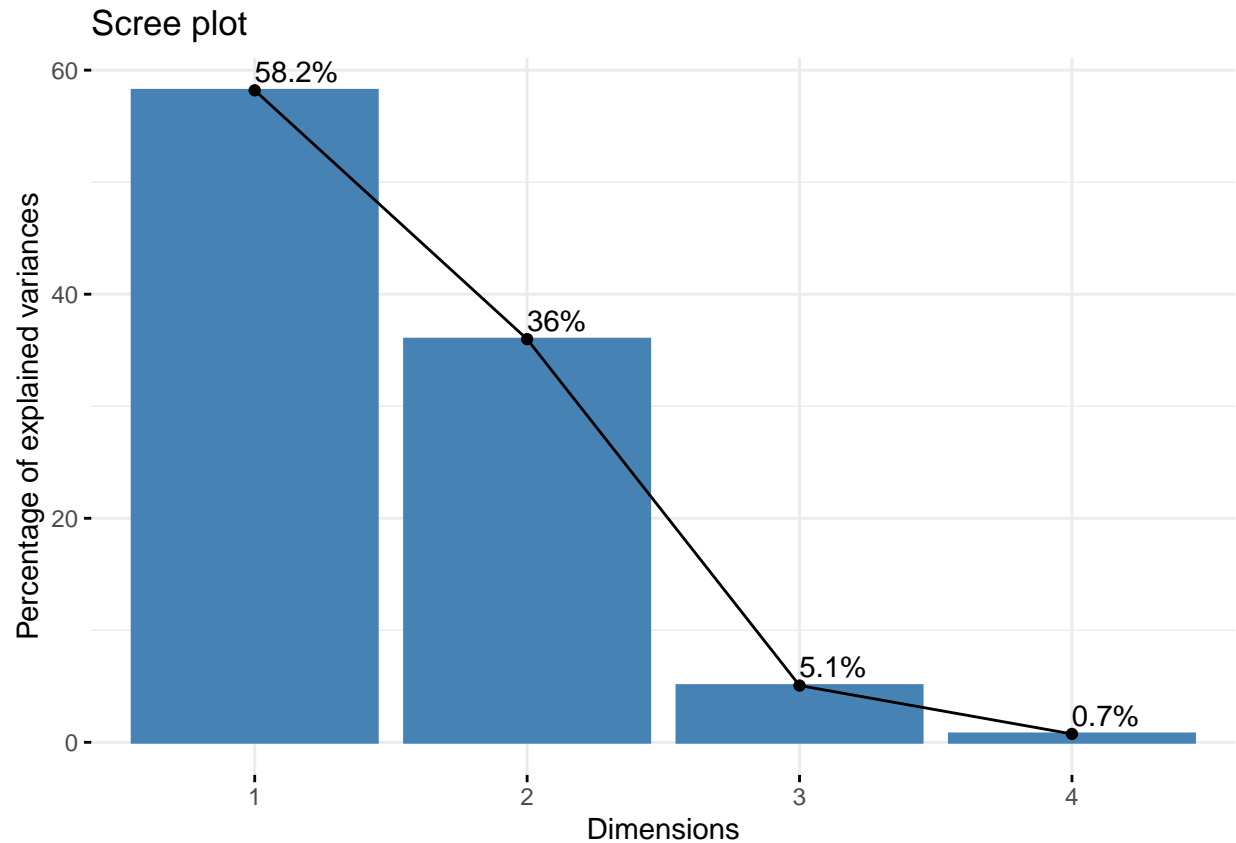
```
I=sum(AFC$eig[,1]);I
```

```
## [1] 0.129594
```

```
AFC$eig
```

```
##          eigenvalue percentage of variance cumulative percentage of variance
## dim 1 0.0754222232          58.1988472          58.19885
## dim 2 0.0466387015          35.9883141          94.18716
## dim 3 0.0065659397           5.0665455          99.25371
## dim 4 0.0009671514           0.7462932         100.00000
```

```
fviz_eig(AFC,addlabels=T)
```



G - Sorties relatives aux points-lignes

```
AFC$call$marge.row
```

```
##      AIX-MARSEILLE      AMIENS      BESANCON      BORDEAUX
##      0.037270072      0.037976309      0.023049019      0.053609836
## CLERMONT-FERRAND      CORSE      CRETEIL      DIJON
##      0.028570511      0.005553594      0.046804276      0.031620173
##      GRENOBLE      LILLE      LIMOGES      LYON
##      0.062630413      0.066643125      0.012840679      0.049629225
##      MONTPELLIER      NANCY-METZ      NANTES      NICE
##      0.041379089      0.040512343      0.070270617      0.022471189
##      NORMANDIE      ORLEANS-TOURS      PARIS      POITIERS
##      0.056338480      0.041860614      0.014670476      0.031299156
##      REIMS      RENNES      STRASBOURG      TOULOUSE
##      0.023594748      0.062598311      0.025296138      0.053898751
##      VERSAILLES
##      0.059612854
```

```
axe1.r=cbind(C1=AFC$row$coord[,1],CTR1=AFC$row$contrib[,1],QLT1=AFC$row$cos2[,1])
round(axe1.r,3)
```

##		C1	CTR1	QLT1
## AIX-MARSEILLE	-0.134	0.887	0.327	
## AMIENS	0.267	3.585	0.387	
## BESANCON	0.195	1.157	0.618	
## BORDEAUX	0.065	0.301	0.106	
## CLERMONT-FERRAND	0.206	1.615	0.481	
## CORSE	0.318	0.743	0.580	
## CRETEIL	-0.128	1.010	0.099	
## DIJON	0.218	1.988	0.862	
## GRENOBLE	-0.213	3.771	0.404	
## LILLE	0.097	0.840	0.035	
## LIMOGES	0.352	2.108	0.762	
## LYON	-0.246	3.974	0.912	
## MONTPELLIER	0.126	0.877	0.699	
## NANCY-METZ	0.208	2.315	0.774	
## NANTES	0.117	1.280	0.334	
## NICE	-0.136	0.550	0.538	
## NORMANDIE	0.234	4.081	0.954	
## ORLEANS-TOURS	0.110	0.677	0.867	
## PARIS	-0.985	18.886	0.949	
## POITIERS	0.254	2.676	0.621	
## REIMS	0.302	2.844	0.649	
## RENNES	0.085	0.595	0.125	
## STRASBOURG	-0.171	0.979	0.447	
## TOULOUSE	-0.034	0.082	0.020	
## VERSAILLES	-0.731	42.179	0.962	

```
axe2.r=cbind(C2=AFC$row$coord[,2],CTR2=AFC$row$contrib[,2],QLT2=AFC$row$cos2[,2])
round(axe2.r,3)
```

##		C2	CTR2	QLT2
## AIX-MARSEILLE	0.154	1.889	0.431	
## AMIENS	0.335	9.126	0.609	
## BESANCON	-0.126	0.785	0.259	
## BORDEAUX	-0.175	3.536	0.767	
## CLERMONT-FERRAND	-0.213	2.772	0.511	
## CORSE	0.209	0.520	0.251	
## CRETEIL	0.373	13.935	0.847	
## DIJON	-0.084	0.474	0.127	
## GRENOBLE	-0.216	6.247	0.414	
## LILLE	0.511	37.371	0.965	
## LIMOGES	-0.171	0.809	0.181	
## LYON	-0.031	0.100	0.014	
## MONTPELLIER	-0.025	0.056	0.028	
## NANCY-METZ	0.100	0.872	0.180	
## NANTES	-0.150	3.375	0.544	
## NICE	-0.030	0.043	0.026	
## NORMANDIE	0.003	0.001	0.000	
## ORLEANS-TOURS	-0.028	0.069	0.055	
## PARIS	-0.050	0.079	0.002	
## POITIERS	-0.129	1.125	0.161	
## REIMS	0.210	2.226	0.314	
## RENNES	-0.221	6.541	0.849	
## STRASBOURG	-0.128	0.889	0.251	

```
## TOULOUSE      -0.229  6.064 0.928
## VERSAILLES    0.093   1.097 0.015
```

H - Sorties relatives aux points-colonnes

```
AFC$call$marge.col
```

```
## [54.2,91.3[ [91.3,99.9[ [99.9,107[ [107,116[ [116,156]
## 0.1999936 0.1982922 0.1993515 0.2021765 0.2001862
```

```
axe1.c=cbind(C1=AFC$col$coord[,1],CTR1=AFC$col$contrib[,1],QLT1=AFC$col$cos2[,1])
round(axe1.c,3)
```

```
##           C1   CTR1  QLT1
## [54.2,91.3[ 0.084  1.891 0.040
## [91.3,99.9[ 0.249 16.350 0.778
## [99.9,107[  0.175  8.123 0.503
## [107,116[   0.020  0.109 0.009
## [116,156]  -0.526 73.526 0.991
```

```
axe2.c=cbind(C2=AFC$col$coord[,2],CTR2=AFC$col$contrib[,2],QLT2=AFC$col$cos2[,2])
round(axe2.c,3)
```

```
##           C2   CTR2  QLT2
## [54.2,91.3[ 0.414 73.620 0.953
## [91.3,99.9[ -0.053  1.176 0.035
## [99.9,107[ -0.166 11.720 0.449
## [107,116[  -0.175 13.313 0.647
## [116,156]  -0.020  0.171 0.001
```